RNA-guided complex from a bacterial immune system enhances target recognition through seed sequence interactions

Wiedenheft, Blake; van Duijn, Esther; Bultema, Jelle; Waghmare, Sakharam; Zhou, Kaihong; Barendregt, Arjan; Westphal, Wiebke; Heck, Albert; Boekema, Egbert; Dickman, Mark

Published in:
Proceedings of the National Academy of Sciences of the United States of America

DOI:
10.1073/pnas.1102716108

Supporting Information

Wiedenheft et al. 10.1073/pnas.1102716108

SI Materials and Methods

Protein Production and Purification. Genomic DNA isolated from strain 14 of *Pseudomonas aeruginosa* (PA14) was used as the template for PCR amplification of the *csy* genes (1). The csy genes were LIC (ligation independent cloning) cloned into expression vectors with and without His_6-tags, using primers V1 and V2, respectively (Table S2). Csy genes were sequentially subcloned into polycistronic expression vectors. Complete nucleotide sequences for each of the vectors are available on the Berkeley MacroLab website (http://macrolab.berkeley.edu/). A synthetic clustered regularly interspaced short palindromic repeat (CRISPR) locus (driven by a T7 promoter) containing eight repeats and seven identical spacers was synthesized by GENEART (Fig. S2*A*).

Csy proteins and the synthetic CRISPR RNA were coexpressed in Rosetta 2 (DE3) *Escherichia coli* cells that were induced with 0.5 mM isopropyl-β-D-thiogalactopyranoside at $OD_{600} = 0.5$ in overnight cultures grown at 16 °C. Cells from the overnight expression cultures were harvested by centrifugation (10,000 × *g*) for 20 min. The cell pellet was resuspended in lysis buffer (15.5 mM disodium hydrogen phosphate, 4.5 mM sodium-dihydrogen phosphate, 500 mM sodium chloride, 10 mM imidazole, 5% glycerol, 0.01% Triton X-100, 2 mM Tris(2-carboxyethyl) phosphine hydrochloride (TCEP), 0.5 mM phenylmethylsulfonyl fluoride, supplemented with protease inhibitors (Roche) and the slurry was sonicated on ice for 2 min in 10 s bursts. The lysate was clarified by centrifugation (22,000 × *g* for 20 min) and the His_6- fusion peptide was bound in batch to Ni-NTA affinity resin (Qiagen). The complex was eluted from the resin in 50 mL lysis buffer containing 300 mM imidazole. The eluted protein was dialyzed at 4 °C overnight against gel filtration buffer (20 mM Hepes pH7.5, 500 mM KCl, 1 mM TCEP, 5% Glycerol). Tobacco etch virus (TEV) protease was added to remove the N-terminal His_6 tag. The protein was concentrated (Amicon) for further purification on a Superose 6 size-exclusion column (GE Healthcare) in a buffer containing 20 mM Hepes (pH 7.5), 0.1 M KCl, 1 mM TCEP.

HPLC Purification of CRISPR RNA (crRNA). All samples were analyzed by ion-pair reversed-phased-HPLC on an Agilent 1100 HPLC with UV260 nm detector (Agilent) using a DNAsep column 50 mm × 4.6 mm I. D. (Transgenomic). The chromatographic analysis was performed using the following buffer conditions: (A) 0.1 M triethylammonium acetate (TEAA) (pH 7.0) (Fluka); (B) buffer A with 25% liquid chromatography MS grade acetonitrile (v/v) (Fisher). The crRNA was obtained by injecting purified intact Csy complex at 75 °C using a linear gradient starting at 15% buffer B and extending to 60% B in 12.5 min, followed by a linear extension to 100% B over 2 min at a flow rate of 1.0 mL/ min. Analysis of the 3′ terminus was performed by incubating the HPLC-purified crRNA in a final concentration of 0.1 M HCl at 4 °C for 1 h. The samples were concentrated to 5–10 μL in a vacuum concentrator (Eppendorf) prior to electrospray ionization mass spectrometry (ESI-MS) analysis.

ESI-MS Analysis of crRNA. ESI-MS was performed in negative mode using an ultra-high resolution time-of-flight mass spectrometer (maXis) or a high-capacity ion trap Ultra post-translational modification Discovery instrument (both from Bruker Daltonics), coupled to an online capillary liquid chromatography system (Ultimate 3000). RNA separations were performed using a monolithic (Polystyrene-divinylbenzene) capillary column (50 mm × 0.22 mm I.D.). The chromatography was performed using the following buffer conditions: (C) 0.4 M 1,1,1,3,3,3,-Hexafluoro-2-propanol (Sigma–Aldrich) adjusted with triethylamine to pH 7.0 and 0.1 mM TEAA, and (D) buffer C with 50% methanol (v/v) (Fisher). RNA analysis was performed at 50 °C with 20% buffer D, extending to 40% D in 5 min followed by a linear extension to 60% D over 8 min at a flow rate of 2 μL/ min. RNA fragmentation was done using 250 ng crRNA with 1U RNase A or RNaseT1 (Applied Biosystems). The reaction was incubated at 37 °C for 4 h. The oligoribonucleotide mixture was separated on a PepMap C-18 RP capillary column (150 mm × 0.3 mm I.D., Dionex) at 50 °C using a gradient starting at 20% buffer C and extending to 35% D in 20 min at a flow rate of 2 μL/ min. The mass spectrometer was set to select a mass range of 250–1500 m/z and the capillary voltage was kept at −3,650 V. Oligoribonucleotides with −2 to −4 charge states were selected for tandem mass spectrometry using collision-induced dissociation.

Protein and Native Mass Spectrometry. Csy complexes were analyzed in 0.15 M ammonium acetate (pH 8.0) at a protein concentration of 5 μM. This protein preparation was obtained by five sequential concentration and dilution steps at 4 °C using a centrifugal filter with a cut-off of 10 kDa (Millipore). Samples were sprayed from borosilicate glass capillaries and analyzed on a LCT electrospray time-of-flight or modified quadrupole time-of-flight instruments (Waters) adjusted for optimal performance in high mass detection (2, 3). Exact mass measurements of the individual Cas proteins were acquired under denaturing conditions (5% formic acid). Subcomplexes in solution were generated by the addition of 2-propanol to the spray solution to a final concentration of 5% (v/v). Instrument settings were as follows; needle voltage was 1.2 kV, cone voltage was 175 V, source pressure 9 mbar. Xenon was used as the collision gas for tandem mass spectrometric analysis at a pressure of $1.5\,10^{-2}$ mbar. The collision voltage varied between 10–200 V.

Isothermal Titration Calorimetry (ITC). DNA oligonucleotides were purchased from Integrated DNA Technologies (IDT) and purified by gel filtration using the buffer described for the purification of the PaCsy complex. Mature crRNAs were isolated from the Csy complex using the mirVana™ kit (Ambion) and eluted in the same buffer. The extinction coefficient for the PaCsy ribonucleoprotein complex was estimated according to the total absorbance at 280 nm using the following formula ($\varepsilon = [(5{,}690 * \mathrm{W57}) + (1{,}280 * \mathrm{Y66}) + (60 * \mathrm{C12})] + [((15200 * \mathrm{A11}) + (12010 * \mathrm{G18}) + (7050 * \mathrm{C16}) + (9800 * (\mathrm{U15}))/2)]$; $\varepsilon = 731{,}120$).

Experiments were performed using an ITC200 instrument (Microcal LLC) at 25 °C. In each experiment the DNA oligo was loaded into the syringe and titrated into the sample with stir rate of 1,000 rpm. The first dummy injection (0.4 μL) was followed by 20 × 2 μL injections. Each injection was separated by 240 s to obtain a flat baseline between injections. Experimental isotherms were fitted according the thermodynamic parameters ΔH (enthalpy) and K_a (association constant), using Origin 7.0. All curves were fit using a nonlinear least-squares "One Set of Sites" (1 : 1 interactions) model. Baseline fitting and peak boundaries were performed automatically with manual corrections.

Electrophoretic Mobility Shift Assays (EMSA). Binding assays were performed by incubating the Csy complex with <0.5 nM of the 5′ ^{32}P-labeled nucleic acid. Each reaction included 25 mM Hepes pH 7.5, 100 mM KCl, 1 mM TCEP, 1% glycerol, 1 mM $MgCl_2$, 1 mg/mL tRNA. All reactions were incubated for 15 min at 37 °C

prior to electrophoresis on 6% polyacrylamide gels. Gels were dried and exposed using phosphor storage screens, scanned with a phosphorimager (GE Healthcare) and quantified using Kaleidagraph (Synergy software).

Electron Microscopy. The Csy complex was stained with 2% uranyl acetate on glow-discharged carbon-coated copper grids. Electron microscopy was performed on a Philips CM120 equipped with a LaB6 tip operating at 120 kV. Images were recorded with a Gatan 4000 SP 4K slow-scan CCD camera at 130,000 magnification at a pixel size of 0.23 nm (after binning the images) at the specimen level with GRACE software for semiautomated specimen selection and data acquisition (4). Single particle projections were selected from micrographs mainly by reference-based automated particle selection procedure incorporated into GRIP (GRoningen Image Processing) software (5). Approximately 28,000 single particles were extracted from 578 electron micrographs. Single particle datasets were analyzed with the GRIP software using multireference alignments and no-reference alignments, multivariate statistical analysis, and hierarchical ascendant classification. The final two-dimensional projection maps were calculated from the best-resolved classes by summing the projections based on the correlation coefficient determined in the alignment step. The number of particle in each class average displayed in each panel (Fig S4).

Small Angle X-ray Scattering. SAXS data were measured at the Advanced Light Source (Lawrence Berkeley National Laboratory) on beamline 7.3.3. Solution scattering of the Csy complex was collected at room temperature (approximately 22 °C) in a 20 μL sample cell at 10 keV (1.24 Å). The sample-to-detector distance was set to 3,056.69 mm resulting in scattering vectors (q) ranging from 0.0141 $Å^{-1}$ to 0.1369 $Å^{-1}$. Two-dimensional scatter curves were transformed and distance distribution functions $P(r)$ were calculated using GNOM (6). Twenty independent bead models were generated by simulated annealing using DAMMIF (7) and these models were aligned, filtered and averaged based on occupancy using DAMAVER (8). The SAXS bead models were converted to volumetric format using the pdb2vol convolution kernel in the Situs software package (9).

1. Lee DG, et al. (2006) Genomic analysis reveals that Pseudomonas aeruginosa virulence is combinatorial. *Genome Biol* 7:R90.
2. Tahallah N, Pinkse M, Maier CS, Heck AJ (2001) The effect of the source pressure on the abundance of ions of noncovalent protein assemblies in an electrospray ionization orthogonal time-of-flight instrument. *Rapid Commun Mass Spectrom* 15:596–601.
3. van den Heuvel RH, et al. (2006) Improving the performance of a quadrupole time-of-flight instrument for macromolecular mass spectrometry. *Anal Chem* 78:7473–7483.
4. Oostergetel GT, Keegstra W, Brisson A (1998) Automation of specimen selection and data acquisition for protein electron crystallography. *Ultramicroscopy* 74:47–59.
5. van Heel M, et al. (2000) Single-particle electron cryomicroscopy: towards atomic resolution. *Q Rev Biophys* 33:307–369.
6. Svergun DI (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J Appl Crystallogr* 25:495–503.
7. Franke D, Svergun DI (2009) DAMMIF, a program for rapid ab initio shape determination in small-angle scattering. *J Appl Crystallogr* 42:342–346.
8. Volkov VV, Svergun DI (2003) Uniqueness of ab initio shape determination in small-angle scattering. *J Appl Crystallogr* 36:860–864.
9. Wriggers W, Milligan RA, McCammon JA (1999) Situs: A package for docking crystal structures into low-resolution maps from electron microscopy. *J Struct Biol* 125: 185–195.
10. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215:403–410.
11. Mojica FJ, Diez-Villasenor C, Garcia-Martinez J, Almendros C (2009) Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology* 155:733–740.
12. Brouns SJ, et al. (2008) Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* 321:960–964.
13. Jore MM, et al. (2010) Structural basis for CRISPR RNA-guided DNA recognition by Cascade. *Nat Struct Mol Biol*, in press.
14. Haurwitz RE, Jinek M, Wiedenheft B, Zhou K, Doudna JA (2010) Sequence- and structure-specific RNA processing by a CRISPR endonuclease. *Science* 329:1355–1358.
15. Ebihara A, et al. (2006) Crystal structure of hypothetical protein TTHB192 from *Thermus thermophilus* HB8 reveals a new protein family with an RNA recognition motif-like domain. *Protein Sci* 15:1494–1499.
16. Moretti S, et al. (2007) The M-Coffee web server: a meta-method for computing multiple sequence alignments by combining alternative alignment methods. *Nucleic Acids Res* 35:W645–648.
17. Gouet P, Courcelle E, Stuart DI, Metoz F (1999) ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics* 15:305–308.

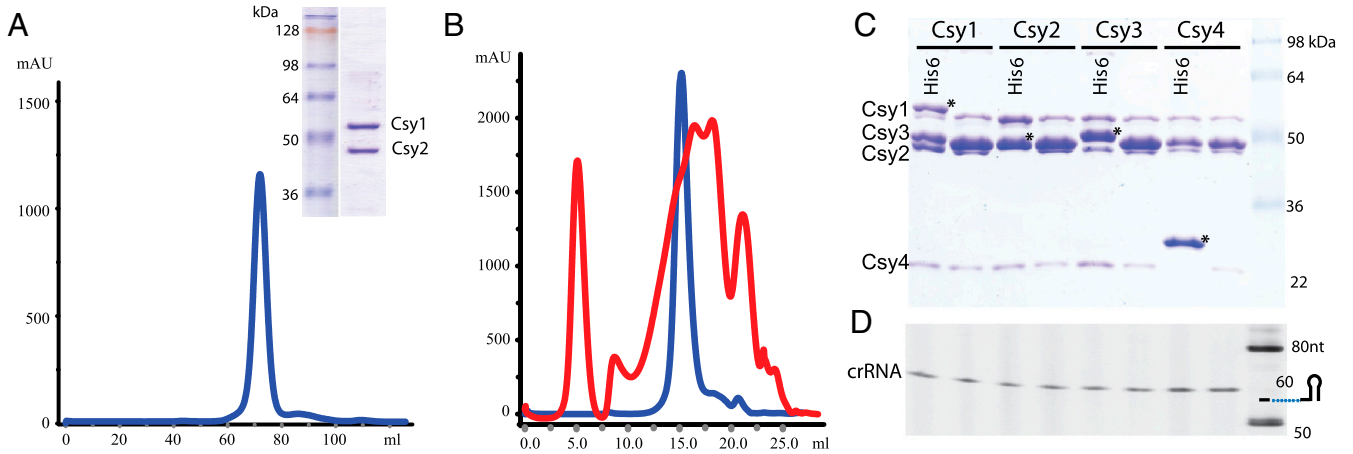
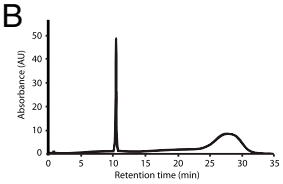


Fig. S1. Purification of the Csy complex. (*A*) Csy1 and Csy2 form a stable heterodimer that elutes as a single peak on a Superdex 200 gel filtration column. The proteins are visible on a Coomassie blue-stained SDS-PAGE (*Insert*). (*B*) Coexpression of the Csy proteins with a CRISPR RNA results in a stable complex that elutes from a calibrated superose 6 column with a retention volume consistent with a molecular mass of approximately 350 kDa (blue line). In contrast the Csy proteins only form transient interactions in the absence of the CRISPR RNA (red line). (*C*) Coomassie blue-stained SDS-polyacrylamide gel of the affinity purified Csy complexes. N-terminal His-tags (*) on Csy1, Csy2, Csy3, or Csy4 can be used as bait to pull down the other untagged Csy proteins. In each case the His-tagged subunit is overrepresented after the first Ni-elution. The His-tag is removed by treatment with the TEV protease, followed by ortho-nickel affinity purification and gel filtration. (*D*) Denaturing polyacrylamide gel electrophoresis of phenol-extracted crRNAs isolated from each of the Csy complexes. The N-terminal His-tag does not interfere with CRISPR RNA processing or with particle assembly.

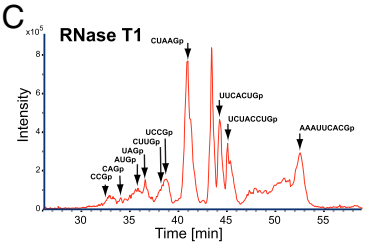
A

TAATACGACTCACTATAGGGCGAATTGAAGGAAGGCCGTCAAGGCCGCATGAGCTCGC-
GACTCCTGCATTAGGAAATTAATACGACTCACTATAGGGTTCACTGCCGTGTAGGCAGC
TAAGAAATTCACGGCGGGCTTGATGTCCGCGTCTACCTGGTTCACTGCCGTGTAGGCAG
CTAAGAAATTCACGGCGGGCTTGATGTCCGCGTCTACCTGGTTCACTGCCGTGTAGGCA
GCTAAGAAATTCACGGCGGGCTTGATGTCCGCGTCTACCTGGTTCACTGCCGTGTAGGC
AGCTAAGAAATTCACGGCGGGCTTGATGTCCGCGTCTACCTGGTTCACTGCCGTGTAGG
CAGCTAAGAAATTCACGGCGGGCTTGATGTCCGCGTCTACCTGGTTCACTGCCGTGTAG
GCAGCTAAGAAATTCACGGCGGGCTTGATGTCCGCGTCTACCTGGTTCACTGCCGTGTA
GGCAGCTAAGAAATTCACGGCGGGCTTGATGTCCGCGTCTACCTGGTTCACTGCCGTGT
AGGCAGCTAAGAAA

B

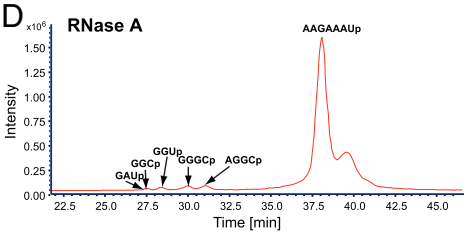


C



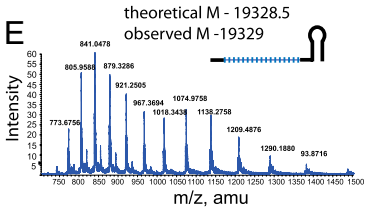
Position	Theoretical mass	Experimental mass	Sequence
C1:G5	1632.995	1632.8	CUAAGp
A6:G14	2902.766	2902.4	AAAUUCACGp
C20:G23	1280.745	1280.2	CUUGp
A24:G26	998.602	998.3	AUGp
U27:G30	1279.761	1279.2	UCCGp
U33:G40	2526.491	2526.4	UCUACCUGp
U42:G48	2221.308	2221.3	UUCACUGp
C49:G51	973.592	973.3	CCGp
U54:G56	998.602	998.3	UAGp
C58:G60	997.617	997.3	CAGp

D



Position	Theoretical mass	Experimental mass	Sequence
A3:U9	2315.438	2315.2	AAGAAAUp
G14:C16	1013.617	1013.2	GGCp
G17:C20	1358.826	1358.2	GGGCp
G23:U25	998.602	998.2	GAUp
G40:U42	1014.602	1014.2	GGUp
A55:C58	1342.826	1342.2	AGGCp
A59:G60	692.433	692.2	AGp

E



F

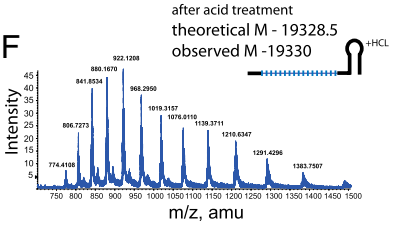


Fig. S2. Mass spectrometry of the CRISPR RNA. (*A*) Sequence of the synthetic CRISPR consisting of eight direct repeats (yellow) and seven identical spacer sequences. CRISPR RNA expression is under control of a T7 promoter (bold underline). (*B*) Ion-pair reversed-phase HPLC purification of mature crRNAs isolated from the Csy complex at 75 °C. (*C*) Base peak chromatogram of the mature crRNA following RNase T1 digestion. RNase T1 cleaves single-stranded RNA 3' of guanine residues. Predominant oligoribonucleotide peaks assigned to the mature crRNA are highlighted (arrows) and a summary of the identified oligos is presented in the table below. (*D*) A similar analysis performed with RNase A and masses for each of the oligos are presented in the table below. (*E*) Multiple charged ESI-MS spectra of the purified mature crRNA before acid treatment and (*F*) after acid treatment. The mass of the mature crRNA is unaffected by acid treatment.

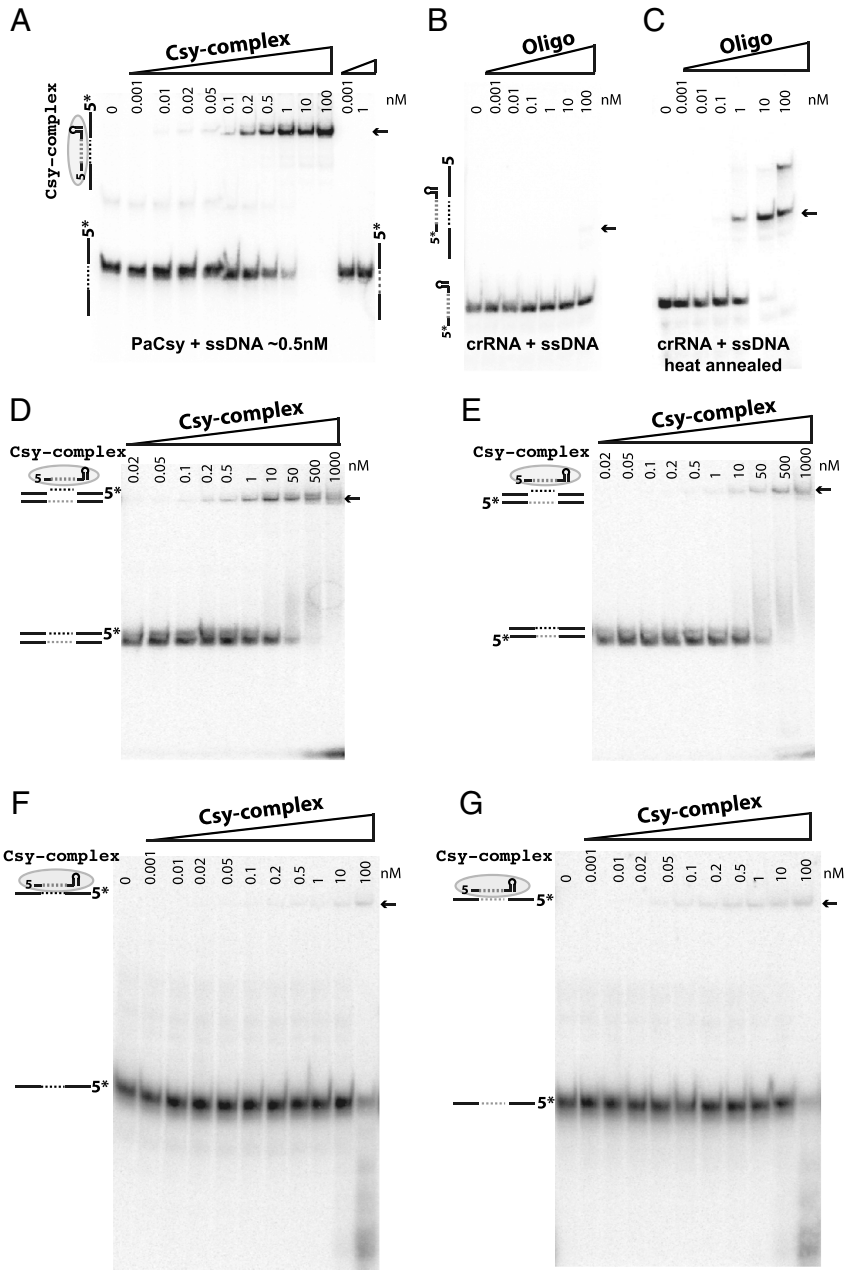


Fig. S3. The Csy complex facilitates target strand-specific hybridization. *A*) A 32-nt target sequence (dashed line) located in the center of the of a 100-nt single-stranded DNA (ssDNA) is recognized and bound by the Csy complex (gray oval) with a K_d of 0.5 nM (left lanes), whereas the Csy complex does not interact with nontarget DNA (far right two lanes). Nucleic acids are 5′ end-labeled with ^{32}P (5*). (*B*) Hybridization between the crRNA and the ssDNA target are modest in the absence of the Csy proteins. A weak interaction between the isolated crRNA and target DNA are detectable only at the highest ssDNA concentration (faint band). (*C*) Low nanomolar binding affinities (approximately 1.2 nM) between the crRNA and the target could be restored in the absence of the Csy complex by heat annealing the pair at 90 °C for 5 min. (*D*, *E*) To test for Csy-mediated strand displacement of dsDNA with a target sequence, we annealed the two complementary ssDNAs used in Fig 1*A* (black dashed target sequence is anneal to the gray nontarget sequence). (*D*) When the target strand in this duplex is 5′ ^{32}P-labeled we observed a doublet in the bound fraction, corresponding to the Csy complex bound to nonduplexed target ssDNA (lower band) and Csy complex bound to the dsDNA duplex (upper band). (*E*) Labeling the nontarget stand results in a single band shift at high protein concentrations. (*F, G*) The Csy complex interacts with long (83-nt) target and nontarget ssRNA with equally low affinity. Secondary structure of long RNAs may limit target site accessibility. Nonspecific degradation of DNA and RNA substrates is observed when protein concentrations exceed 1,000 nM. All reactions were incubated at 37 °C for 15 min prior to electrophoresis except the reactions in panel C, which were done at 90 °C for 5 min. Gels in panels A, B and E are shown in Fig. 2 of the main text.

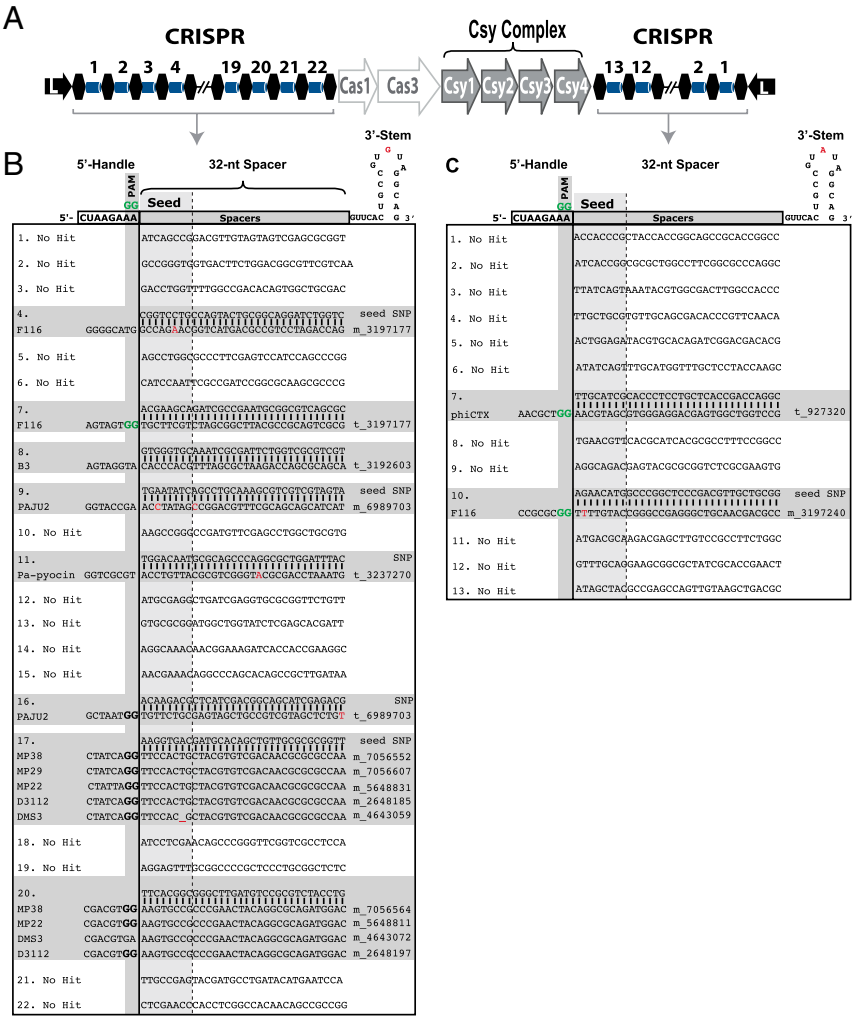


Fig. S4. Spacers from both CRISPR loci map to phage known to infect *Pseudomonas*. (*A*) Two CRISPR loci flank a cassette of *cas* genes in the *P. aeruginosa* (PA14) genome. Each CRISPR consists of a series of direct repeats (black hexagons) that are separated by unique spacer sequences (blue cylinders). Both CRISPRs are flanked by a leader sequence ("L" in black arrow). (*B*, *C*) A schematic of the mature crRNA from each CRISPR loci is shown above the two tables that list all the spacer sequences form that CRISPR. The repeat sequence within each CRISPR is highly conserved and the repeat sequence between the two CRISPRs differs by only a single nucleotide (*G* vs. *A*) in the loop of the crRNA (red). Each spacer was queried against the nonredundant database [National Center for Biotechnology Information (NCBI)] using BLAST (10). Whereas most spacers are unique ("No Hit"), the nonunique spacers usually map to phage known to infect *Pseudomonas*. Sequences complementary to spacers have been termed; proto-spacers and all identified spacer/proto-spacers pairs are aligned (gray boxes). A proto-spacer adjacent motif (PAM) has been previously identified (11). Sequence 3′ of the proto-spacer is shown for all hits and PAMs are shown in green. The identification of each target is listed below the spacer number (*Left*) and the target strand (mRNA or template), as well as the NCBI gene identification numbers are listed on the right (i.e., m_3197177). Single nucleotide polymorphisms are shown in red.

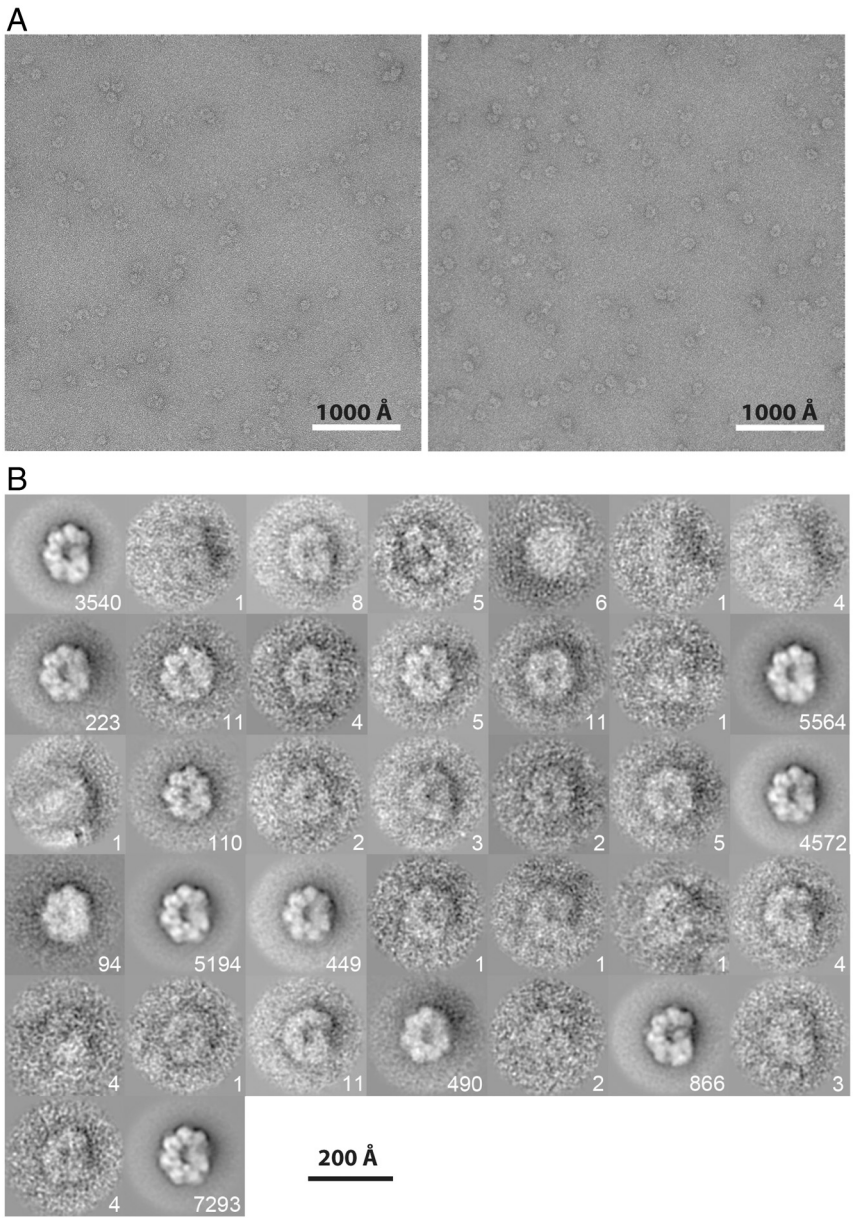


Fig. S5. Transmission electron microscopy of the PaCsy complex. (*A*) Two representative images recorded at 130,000 times magnification. (*B*) Class averages after hierarchical classification from 28,452 particles into 37 classes. The number of particles in each class is displayed in the lower right corner of each image.

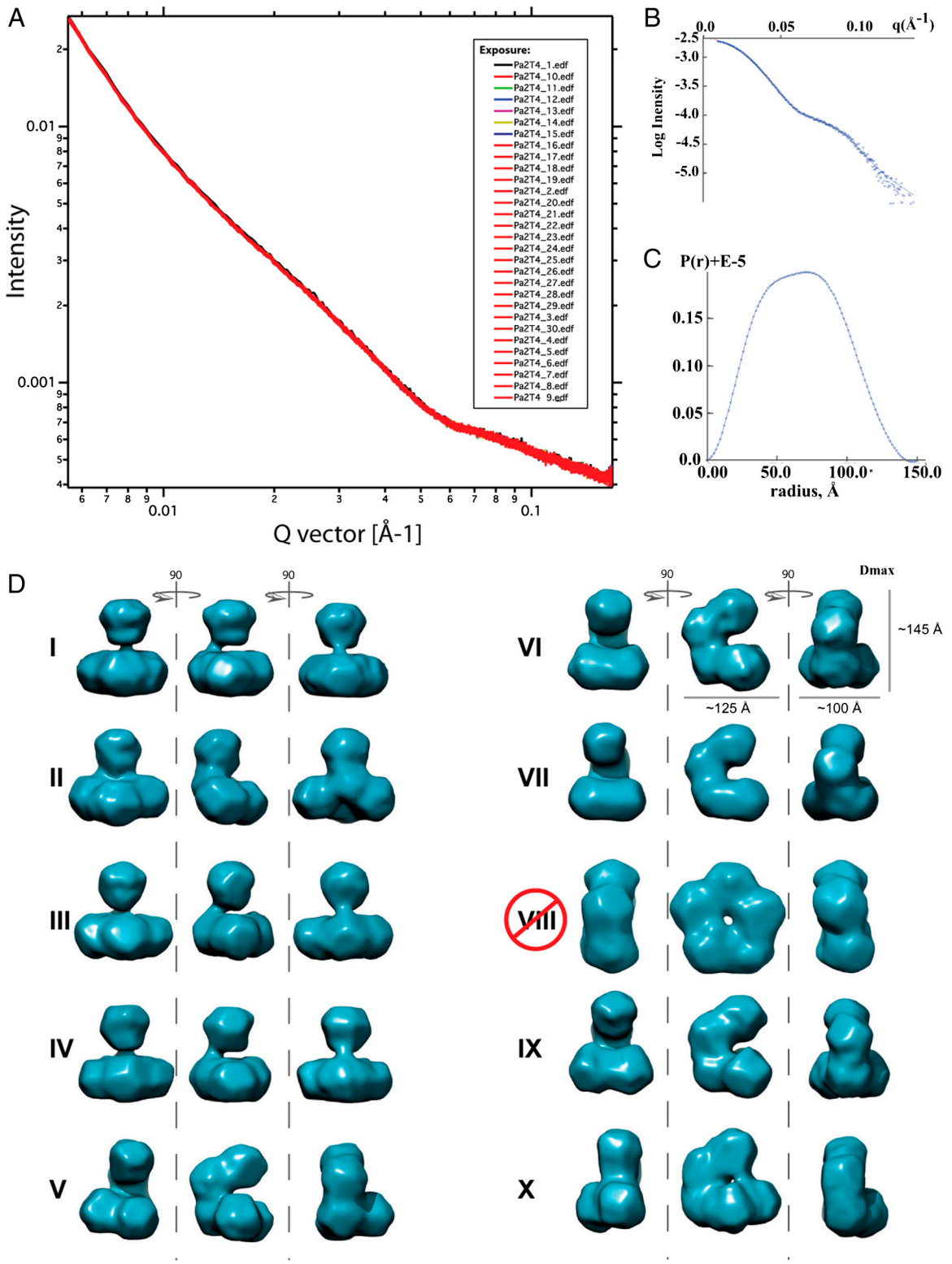


Fig. S6. SAXS scattering curves are superimposable. (*A*) Two-dimensional scattering curves generated from thirty exposures are superimposable and linear in the Guinier region. (*B*) These data were radially integrated, averaged and background subtracted, resulting in a final scatter plot that includes scattering vectors (*q*), ranging from 0.0141 to 0.1369 $Å^{-1}$. (*C*) The paired-set of distances (*Pr*) between scattering electrons indicates that the most frequently sampled interatomic distance in the Csy complex is 52.0 Å. (*D*) Ten of the original twenty bead models were selected at random (I-X) and each model is shown in three orientations. This montage shows a general open-ring shape that is recovered in the individual reconstructions. Model VIII (Ø) did not meet the filtering criteria (normalized spatial discrepancy >Mean + 2 * Variation) and was not included in the averaged model.

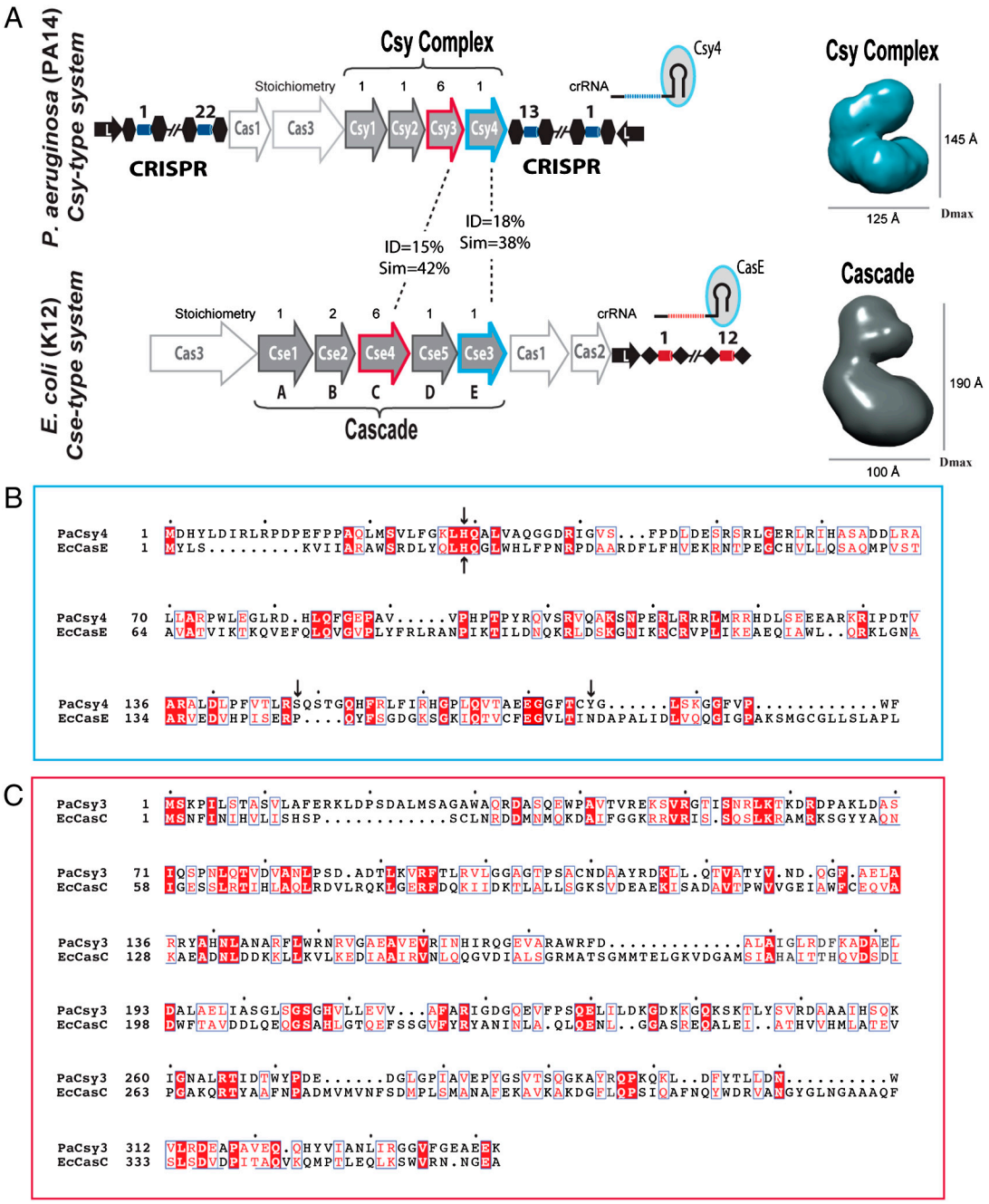


Fig. S7. Comparison of the Csy and Cse-type CRISPR-mediated immune systems. (*A*) Genome organization of the Csy-type immune system from *P. aeruginosa* (PA14) and the Cse-type immune system from *E. coli* (K12). The "core" *cas* genes are shown as white arrows and the subtype specific cas genes are display as gray arrows. Csy4 and Cse3 (CasE) are CRISPR specific endoribonucleases that exclusively bind and specifically cleave their own CRISPR RNAs (gray oval with blue outline). The Csy proteins ($Csy1_1:Csy2_1:Csy3_6:Csy4_1$) and a crRNA assemble into a 350 kDa ribonucleoprotein complex. The three dimensional shape of this complex has been determined by small angle X-ray scattering (SAXS model displayed far right). This complex is similar to, but distinct from the Cascade complex (CRISPR-associated complex for antiviral defense; gray) from *E. coli* (12, 13). Cascade is a 405 kDa complex comprised of an unequal stoichiometry of Cse proteins, which have been renamed CasA-D ($CasA_1:B_2:C_6:D_1:E_1$) and a crRNA. A SAXS reconstruction of this complex reveals a seahorse shaped particle with a twist along the spine (gray) (13). (*B*) Amino acid sequences of Csy4 from the PA14 strain of *P. aeruginosa* (PaCsy4) and Cse3 form *E. coli* (EcCasE) are 18% identical. Structure guided mutations of Csy4 have identified key catalytic residues (H29, S148 and Y176, arrows) (PDB ID: 2xli, 2xlj, and 2xlk) (14). A structure of Cse3, a CasE homolog from *Thermus thermophilus* (PDB ID: 1WJ9), has also been published and mutational analyses in *E. coli* have identified H20 as a catalytically essential residue (arrow) (12, 15). (*C*) The Csy complex contains 6 copies of Csy3, whereas Cascade contains 6 copies of Cse4 (CasC). Although these two sequences share limited sequence identity (15%), the stoichiometry of these proteins indicates that they may contribute a common architectural role to their respective complexes. Sequence alignments were preformed using M-Coffee (16) and display using ESPript (17).

Table S1. Exact masses of the individual Csy proteins determined by denaturing mass spectrometry

	Stoichiometry (Csy1:Csy2:Csy3:Csy4:crRNA)	Theoretical mass (Da)	Experimental mass (Da)
Denaturing mass spectrometry data			
Csy1	1:0:0:0:0	48,998.1^{-M}	49,000.9 ± 2.0
Csy2	0:1:0:0:0	36,070.0^{-M}	36,071.7 ± 1.7
Csy3	0:0:1:0:0	37,399.2^{-M}	37,401.0 ± 1.2
Csy4 + tag	0:0:0:1:0	23,583.8	23,585.0 ± 0.9
Csy4	0:0:0:1:0	21,532.6	21,533.8 ± 1.9
Native mass spectrometry data			
Complex	1:1:6:1:1	350,324	350,414 ± 22
Complex (Csy4 still tagged)	1:1:6:1:1	352,376	352,494 ± 16
Complex-Csy1/2	0:0:6:1:1	265,256	265,307 ± 9

Masses for the Csy complexes observed under native mass spectrometry conditions are also listed. To calculate the theoretical mass of the complexes a mass of 19,328.5 Da was used for crRNA.

Table S2. Primer sequences used for ligation independent cloning of the *csy* genes

Gene	PA_Gene #	Gene ID	Primer	Primer sequence*
Csy1	PA_33310	4380506	C1 F v1	TACTTCCAATCCAATGCAACCTCTCCCCTCCCAAC
			C1 F v2	TTTAAGAAGGAGATATAGATCATGACCTCTCCCCTCCC
			C1 rev v1	TTATCCACTTCCAATGTTATTATCAGTCACGCTCATCTTCGA
			C1 rev v2	TTATGGAGTTGGGATCTTATTATCAGTCACGCTCATCTTCGA
Csy2	PA_33320	4380507	C2 F v1	TACTTCCAATCCAATGCAAGCGTGACTGATCCCGAG
			C2 F v2	TTTAAGAAGGAGATATAGATCATGAGCGTGACTGATCCC
			C2 rev v1	TTATCCACTTCCAATGTTATTATTATGCGATGGCGTGTTCGA
			C2 rev v2	TTATGGAGTTGGGATCTTATTATTATGCGATGGCGTGTTCGA
Csy3	PA_33330	4380506	C3 F v1	TACTTCCAATCCAATGCATCCAAGCCAATACTGAGCAC
			C3 F v2	TTTAAGAAGGAGATATAGATCATGTCCAAGCCAATACTGAGC
			C3 rev v1	TTATCCACTTCCAATGTTATTATTACTTCTCTTCGGCTTCACC
			C3 rev v2	TTATGGAGTTGGGATCTTATTATTACTTCTCTTCGGCTTCACC
Csy4	PA_33340	4380486	C4 F v1	TACTTCCAATCCAATGCAGACCACTACCTCGACATTC
			C4 F v2	TTTAAGAAGGAGATATAGATCATGGACCACTACCTCGACAT
			C4 rev v1	TTATCCACTTCCAATGTTATTATCAGAACCAGGGAACGAAAC
			C4 rev v2	TTATGGAGTTGGGATCTTATTATCAGAACCAGGGAACGAAAC

*The underlined sequence for each primer corresponds to the LIC tag.